Gender Differences in Connectome-based Predictions of Individualized Intelligence Quotient and Sub-domain Scores

SUPPLEMENTARY FILES

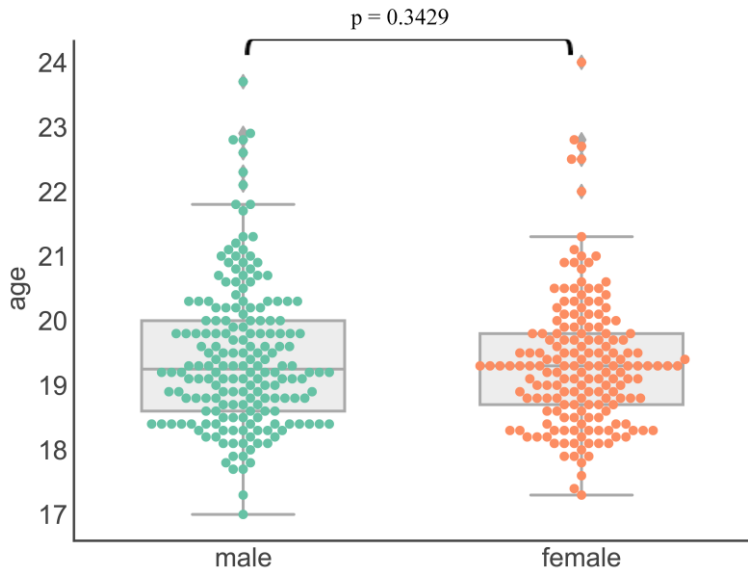

Figure S1. Males and females did not differ in age. A total of 360 subjects (174 F/186 M, 19.4 ± 1.1 years, in range of 17-24 years) were included in UESTC data set. Specifically, males have a mean age of 19.46 ±1.14 years, while females have a mean age of 19.35 ± 1.04 years. Two-sample t-test revealed that there is no significant difference between them in age (t[358] = 0.9498, p = 0.3429).

S1: COBRE dataset.

A total of 178 participants with complete imaging data and matching psychological assessment were recruited as part of a multimodal schizophrenia center for biomedical research excellence (COBRE) study at the Mind Research Network (Sui J et al. 2013) (http://cobre.mrn.org). Due to the fact that UESTC were all college students which means a higher education level and IQ scores ($p = 8.4\times10^{-4}$) relative to COBRE individuals, 52 Hispanic or Latino subjects were excluded. Six individuals with excessive head motion, defined a priori as >2 mm translation, or >3° rotation during the run, were also excluded. Finally, 120 subjects (42 females/78 males; 18 – 65 years, mean age = 38.0 years) were retained for further analysis, in which 9 were diagnosed with bipolar disorder (BP), 51 with schizophrenia or schizoaffective disorder, and 60 HCs. Females and males are age- (p = 0.52) and education- (p = 0.61) matched. Participants provided written informed consent according to institutional guidelines required by the Institutional Review Board at the

University of New Mexico (UNM), and were paid for their participation. All subjects were screened and excluded if they had history of neurological disorder, history of mental retardation, history of severe head trauma with more than 5 min loss of consciousness, or history of substance abuse, or dependence within the last 12 months (except for nicotine). HCs were free from any Axis I disorder, as assessed with the SCID-NP (Structured Clinical Interview for DSM-IV-TR, Non-patient version). Schizophrenia or bipolar disorder was diagnosed according to DSM-IV–TR criteria on the basis of a structured clinical interview (First MB, Spitzer, R.L., Gibbon, M. and Williams, J.B. 1995). All patients were on stable medication prior to the fMRI scan session. Symptom scores were determined based on the positive and negative syndrome scale (PANSS).

S2: Data acquisition and preprocessing

HCP dataset. Data acquisition has been described in detail elsewhere (Ugurbil K et al. 2013). In this study, we used the same main descriptions for data collections that we used in our previous work (Zuo N et al. 2018). For the sake of completeness, we repeat these main descriptions for data collection here: resting-state fMRI data were collected at Washington University in St. Louis using a 3T Skyra (Siemens, Erlangen, Germany) with a 32-channel head coil. The primary scanning parameters were repetition time (TR), 720 ms; echo time (TE), 33.1 ms; flip angle, 52°; field of view, 208 × 180 mm; slice thickness, 2.0 mm; and voxel size, 2.0 mm isotropic cube (Smith SM et al. 2013).

COBRE dataset. MRI scans were performed on a 3-T Siemens Trio scanner with a 12-channel radio frequency coil at the Mind Research Network. Briefly, resting state data were collected with single-shot full k-space echo-planar imaging (EPI) with ramp sampling correction using the inter commissural line (AC/PC) (anterior commissure/posterior commissure) as a reference (TR = 2 s, TE = 29 ms, matrix size = 64 × 64, flip angle = 75 °, slice thickness = 3.5 mm, slice gap = 1.05 mm, field of view (FOV) 240 mm, matrix size = 64 × 64, voxel size = 3.75 mm × 3.75 mm × 4.55 mm. Resting-state scans were a minimum of 5 min, 4 s in duration (152 volumes). Subjects were instructed to keep their eyes open during the scan and stare passively at a foveally presented fixation cross, as this is suggested to facilitate network delineation compared to eyes-closed conditions and helps ensure that subjects are awake. Detailed data acquisition can be found in (Sui J *et al.* 2013).

The HCP data were already preprocessed, well aligned, and registered to the Montreal Neurological Institute (MNI) 2-mm standard space when we received it. The main preprocessing steps taken included (Glasser MF et al. 2013): (1) gradient nonlinearity distortion; (2) 6 degrees of freedom (DOF) FSL/FLIRT-based motion correction; (3) FSL/top-up-based distortion correction; (4) registration to a T1 space image; and

(5) FSL/FNIRT-based registration to MNI 2-mm space. After receiving the above preprocessed data from HCP, we further band-pass-filtered the data at 0.009–0.08 Hz to reduce low-frequency drift and high-frequency noise (Vatansever D et al. 2015). The mean signal of the white matter and cerebrospinal fluid (CSF) and the movement parameters and its derivatives (in the Movement_parameters.txt file in the HCP S500 release) were regressed out as confounding factors. Preprocessing for the COBRE dataset can be found in (Sui J *et al.* 2013). Briefly, the SPM8 software package was employed to perform fMRI preprocessing. The first four volumes are discarded to remove T1 equilibration effects. Slice timing was performed with the middle slice as the reference frame. Images were realigned using INRIalign, a motion correction algorithm that is unbiased by local signal changes (Freire L et al. 2002). Subsequent preprocessing included spatial normalization into the standard Montreal Neurological Institute (MNI) space (Friston KJ et al. 1995), temporal band-pass filtering (0.01 Hz to 0.08 Hz) and resampling to 3 mm × 3 mm × 3 mm. Before smoothing, we further regress out the six motion parameters for each slice to remove the motion effect. Finally, data were spatially smoothed with a 8 mm Gaussian kernel. Notably, head motion, calculated as mean frame-to frame displacement (which was correlated with IQ scores in the COBRE dataset (r = -0.27, p = 0.002)) was regressed out of the data.

Table S1. Prediction results of IQ-predictive models after controlling for potential confounds

	Main results		Ruling out head motion and age	
	Males	**Female**	**Males**	**Females**
UESTC	r= 0.46, p= 3.1×10^{-12}	r= 0.72, p= 3.2×10^{-29}	r= 0.41, p= 9.7×10^{-9}	r= 0.69, p= 9.8×10^{-27}
HCP	r= 0.253, p= 0.02	r= 0.289, p= 0.001	r= 0.281, p= 0.016	r= 0.289, p= 0.001
COBRE	r= 0.23, p= 0.04	r= 0.40, p= 0.008	r= 0.20, p= 0.08	r= 0.39, p= 0.01

Compared with the discovery dataset UESTC, participants in validation datasets of HCP and COBRE include a wide range of age (UESTC: 17 – 24 years; HCP: 22 – 65 years; COBRE: 18 – 65 years). Although age did not show statistically significant correlations (p>0.2) with intelligence levels, we nonetheless want to minimize potential age confounds in prediction. In addition, in COBRE dataset, head motion (calculated as mean frame-to frame displacement) revealed significant correlation with IQ scores (r = -0.27, p = 0.002). Thus, we further calculated the partial correlations between the predicted and observed IQ scores while adopting age and mean frame-to-frame motion as control measurements. Regarding the discovery dataset, we found that controlling for age and head motion resulted in slightly attenuated correlations. With regard to the validation datasets, results indicate comparable prediction performance after controlling for these potential confounds, except that male-specific predictive model achieved marginally significant correlation for COBRE males.

S3: Individualized Prediction and Test of different Regression models:

We adopted a leave-one-out cross-validation strategy to evaluate the prediction performance, which is also the recommended approach named 'learn-predict separation' to avoid leakage by (Kaufman S et al. 2012), and has been widely employed by most existing prediction studies (Finn ES et al. 2015; Rosenberg MD et al. 2016; Beaty RE et al. 2018; Reggente N et al. 2018; Yamashita M et al. 2018; Yip SW et al. 2019). In the feature selection step, with the ReliefF algorithm (Robnik-Sikonja M and I Kononenko 1997), every training feature was assigned with a weight statistically accounting for its relevance to the IQ scores. By determining top ***m*** weighted features, we can exclude redundant features effectively (Meng X et al. 2017). Note that the optimal number of ***m*** for ReliefF needs to be determined by the data, and once determined, it remains constant across all cross-validation loops. Since we employed a leave-one-out cross-validation approach, a total of *N* (sample size) prediction models were constructed across N loops.

a) For loop 1, we first selected the top *m* important FCs determined by ReliefF and enter these *m* features with the corresponding IQ scores into LASSO model across N-1 training subjects, yielding a prediction model.
b) Then, we extracted the same *m* features from one testing subject and submitted them to the constructed prediction model, generating a predicted IQ score for this subject.
c) Afterwards, we repeated step a)-b) for loop 2, 3…N, until all subjects have a predicted IQ score. Notably, the value for *m* was constant for all N loops.

To obtain the optimal value, *m* was tested ranging from 10 to 1000 with a 20 interval, and then ranging from 1000 to 10000 with a 50 interval. Stated differently, for each of the tested threshold (10:20:1000, 1000:50:10000), we repeated the whole prediction procedure **a)-c)**. Consequently, for each of the tested threshold, we derived a prediction performance, and the *m* value that yielded the highest prediction accuracy was finally determined as the optimal parameter. As displayed in **Figure S3a**, an optimal feature number of *m* = 3600, 3300 and 3900 for males, females and all subjects were determined for the prediction.

Additionally, the feature selection step was adopted to reduce the redundancy, simplify the fitted model, and enhance generalization, given that feature dimension in our study considerably overwhelms the sample size. The assessment of the effectiveness of features selection really depended on the derived prediction performance, since our ultimate goal was to achieve successful prediction of IQ scores with as high as possible prediction accuracy.

Notably, the determination of *m* also follows existing studies (Dosenbach NU et al. 2010; Meng X *et al.* 2017; Greene AS et al. 2018; Liu Z et al. 2018). Generally, all the above studies employed a prediction framework which incorporates a filter-based feature selection method plus a regression algorithm, within a fully cross-validated analysis. In the feature selection step, a subset of whole-brain features were selected under a predefined threshold. And then these selected features were submitted to a regular regression method (LASSO, multiple linear regression, elastic net, support vector regression) to construct a prediction model. The optimal parameter was always tested from a series of candidate values, and the value that yielded the highest prediction accuracy was finally determined as the optimal parameter.

In addition, ReliefF assessed the importance of each feature by calculating the relative distance between a randomly selected subject and its *k* neighbor subjects. More details can be found in (Robnik-Šikonja M and I Kononenko 2003). Therefore, *k* is another hyper parameter that needs to be determined. Generally, if we set *k* to 1, the estimates computed by ReliefF can be unreliable for noisy data. If we set *k* to a big value, ReliefF can fail to find important attributes. Empirically, 10 is the most commonly used value, and it is also the recommended parameter in Matlab's built-in ReliefF function. Consequently, we set *k* to this default value as our previously published papers (Meng X *et al.* 2017; Jiang R et al. 2018). Moreover, we also performed additional analysis to explore the influence of *k* on prediction accuracy for females (Urbanowicz RJ et al. 2018). Instead of being set as a default values of 10, this time, *k* values were tested ranging from 10 to 173 (10, 15, 20:10:80, 100:20:160, 173). Then, for each of the above *k* values, we rean the whole prediction procedure within rigorous cross-validations while ranging *m* from 3000 to 4000 with a 100 interval. Notably, *m* was not tested from 10 to 10000, because from our main results, we found that the highest prediction accuracy was usually achieved between *m*=3000 to 4000. **Supplementary Figure S3b** demonstrates the prediction results. Overall, our prediction pipeline achieved comparable prediction accuracies when setting *k* to 10, 15, or 20, and a larger *k* value yielded slightly attenuated prediction performance. Consequently, *k* exerts little influence on the prediction accuracy for our data.

To compare the prediction performance of different regression algorithms, a total of 4 popular linear regression models were used here, including least absolute shrinkage and selection operator (LASSO), ridge regression, the Elastic net and relevance vector regression (RVR). Specifically, LASSO applies an L1-norm penalty, which minimizes the sum of the absolute regression coefficients. Notably, LASSO achieves a sparse model, where at most *N* features can be selected in the final model (*N* is the sample size) (Tibshirani R 1996). In contrast, the ridge regression applies an L2-norm penalty, which minimizes the sum of the square of the

regression coefficients and retains all features in the model. The elastic net uses a combination of L1-norm and L2-norm penalties through a mixing parameter α. RVR is a sparse kernel method formulated in a Bayesian framework (Zou H and T Hastie 2005). Compared with non-linear models, linear methods can reduce the overfitting problem to some extent, and have good performance in interpreting how different features influence the prediction. Furthermore, to evaluate the effectiveness of feature selection, prediction procedures without ReliefF were also tested for all four regression models, where the whole-brain FCs were used as predictions.

For LASSO, elastic net and ridge regression methods, the regularization penalty is controlled through the regularization parameter λ. For elastic net, the mixing parameter α was chosen from 0.1 to 1.0, with a 0.1 step. In the training phase, the three algorithms were implemented in a 10-fold cross validation, where the value of λ that gives minimum mean cross-validated error was selected by default. Consequently, no parameters need to be tuned manually in this step. No parameters need to be determined for RVR.

For the feature selection method ReliefF, functions provided in the statistics and machine learning toolbox in Matlab were used. Compared with the majority of the feature selection measures, ReliefF is able to effectively select useful features and has a low computational complexity (Stokes ME and S Visweswaran 2012). For regression algorithms LASSO, ridge regression and Elastic net, we made use of the Glmnet package provided by (Friedman J et al. 2010); for RVR, the kernel methods package Kernlab provided by (Karatzoglou A et al. 2004) was employed in R.

Table S2: Prediction results of four regression models and results without feature selection.

Correlation *r*	All subjects	Male subjects	Female subjects	Steiger's z/p-value
ReliefF+LASSO	**0.5122**	**0.4622**	**0.7212**	3.86/0.0001
ReliefF+ridge	0.4787	0.3010	0.4918	2.13/0.033
ReliefF+elastic net	0.4313	0.2787	0.6481	4.53/<0.0001
ReliefF+RVR	0.2189	0.1353	0.2359	0.97/0.332
LASSO	0.3668	0.1678	0.6802	6.28/<0.0001
Ridge	0.4345	0.1815	0.4295	2.61/0.009
Elastic net	0.3449	0.1830	0.6655	5.91/<0.0001
RVR	0.2413	0.0859	0.2609	1.76/0.078

Table S2 demonstrates the prediction results of IQ scores using 4 regression methods. In the prediction of IQ scores, a prediction framework integrating feature selection and regression technique was adopted within a leave-one-out cross-validation (LOOCV) strategy. To compare the prediction performance of different regression algorithms, a total of 4 popular linear regression models were used here, including LASSO, ridge regression, the Elastic net and RVR. Due to the fact that the number of features selected in ReliefF has an important influence on the prediction performance, we performed the LOOCV prediction framework by ranging the feature number applied in ReliefF from 10 to 1000 with a 20 interval, and then ranging from 1000 to 10000 with a 50 interval. Results provided in the table were all derived with the optimal feature number, which can generate the highest correlation *r* between the predicted and true IQ scores. Prediction results using whole-brain FCs without feature selection were also provided. Compared with prediction incorporating feature selection, prediction with only regression technique impaired the prediction performance almost for all four regression methods. Note that female IQ was always more predictable than males regardless of the type of feature selection in all four regression models ($p = 0.0045$). Moreover, a Steiger's z-test for testing differences between two independent correlations (Steiger JH 1980) revealed that the difference in prediction performance for males and females reached significance in most cases ($p < 0.05$).

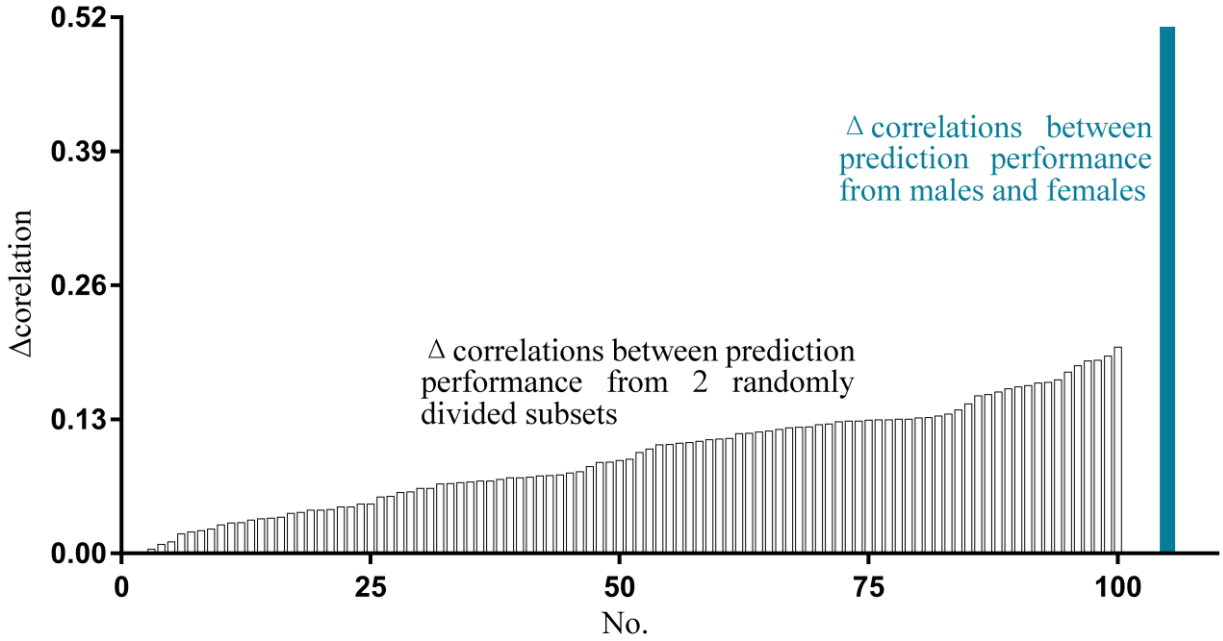

Figure S2. In our main results, female IQ was always more predictable than males regardless of regression models employed. To further validate that the difference in predictability was a reflection of actual gender difference, which was not influenced by data division, we randomly divided the whole sample (n=360) into two halves (n=180) and reran the prediction procedure with each subset within rigorous cross validation. This process was repeated 100 times. Notably, to exclude the influence brought by potential confounds such as algorithm hyper parameter, as well as to reduce computational complexity, we employed the LASSO approach here. In Table S2, we have shown that LASSO achieved a prediction accuracy of r[female]= 0.6802 for females, which is significantly higher (Δ=0.5124) than that for males r[male]= 0.1678 ($p<0.0001$). This figure demonstrates the absolute values of Δcorrelations of the prediction performance derived from the 2 randomly divided halves (black bar), which are sorted in an ascending order. Apparently, the Δcorrelations between the randomly divided subsets (max Δ=0.2005) were much lower than that between males and females (green bar). Consequently, the Δcorrelation of prediction performance between males and females was significantly higher than that between any 2 randomly divided subsets under 100 permutation test ($p<1/(1+100)=0.01$).

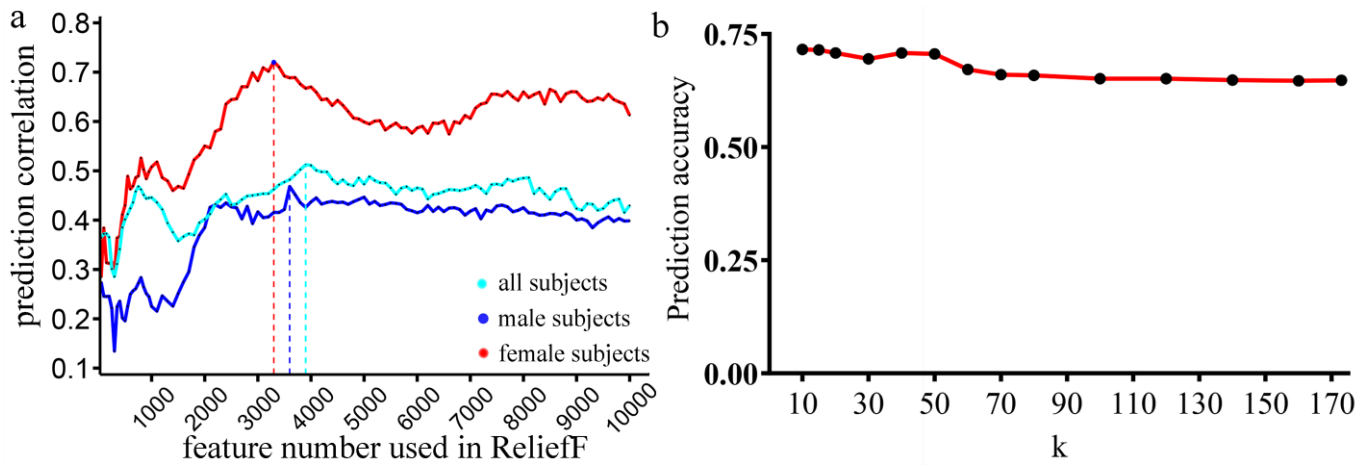

Figure S3. Influence of ReliefF feature number and the number of neighbors on prediction. (a) The number of features selected in ReliefF was tested ranging from 10 to 1000 with a 20 interval, and then ranging from 1000 to 10000 with a 50 interval. Ultimately, an optimal feature number of 3600, 3300 and 3900 for males, females and all subjects were determined in the prediction. Interestingly, females exhibit significantly higher IQ-predictability than males regardless of the feature number used ($p < 0.001$). (b) In addition, ReliefF assessed the importance of each feature by calculating the relative distance between a randomly selected subject and its *k* neighbor subjects. This was performed only for females.Results demonstrated that our prediction pipeline achieved comparable prediction accuracies when setting *k* to 10, 15, or 20, and a larger *k* value yielded slightly attenauted prediction performance.

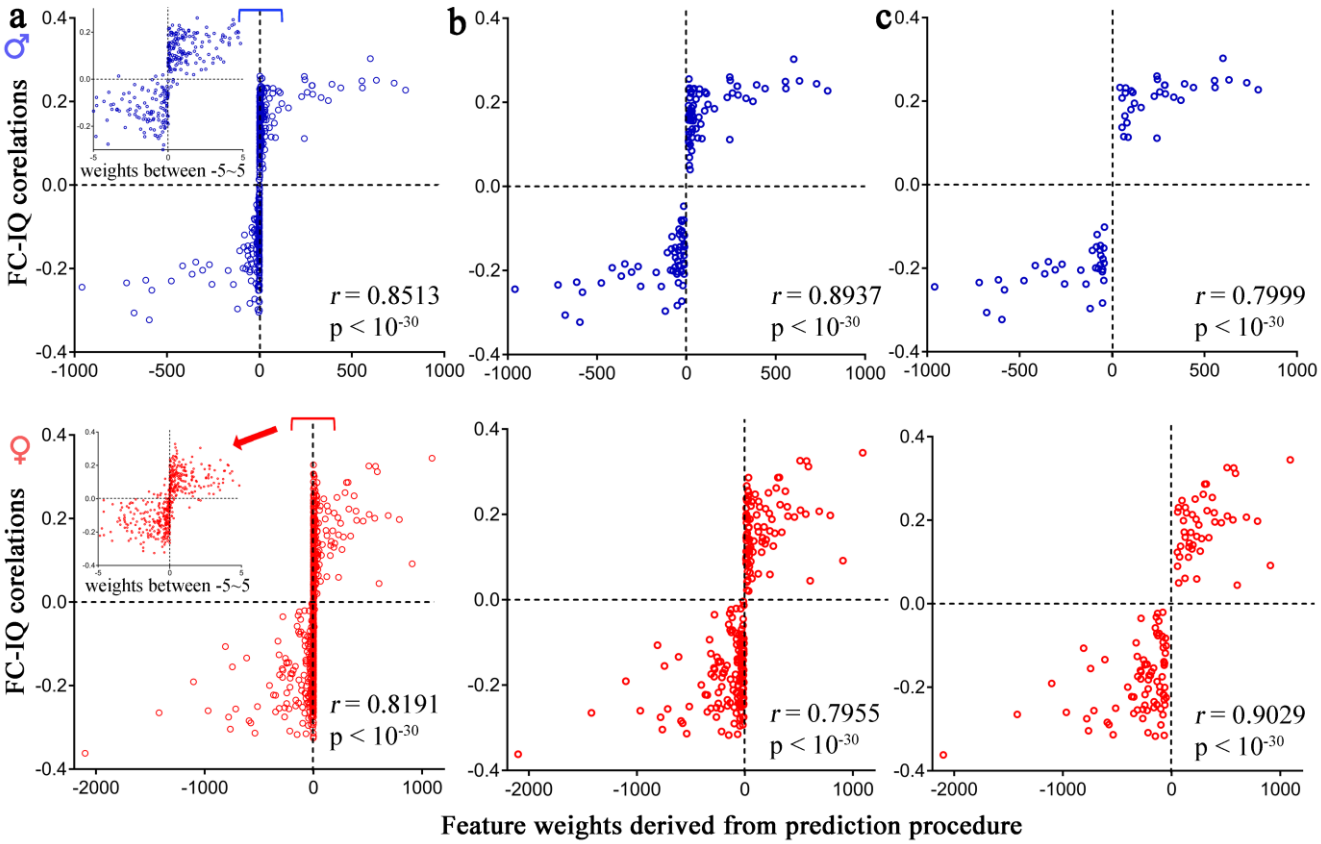

Figure S4. The direction of the identified feature weights relates to the true direction of the correlations between FC and IQ scores. Since we applied a cross-validated prediction strategy to estimate the IQ scores, in each iteration, slightly different FCs were selected. Overall, across all cross-validation loops, a total of 451 and 789 different features were selected for males and females respectively, which were the FCs used to estimate the relative feature importance (**Figure 4a**). We calculated the Person's correlations *R* between each of these selected FCs and IQ scores for males and females respectively. These *R* values represent the true direction of the relationships between functional connectivity and individuals' IQ scores, where a higher absolute value indicates a greater relevance between them. To assess whether the direction of the feature weights calculated from our prediction procedure is consistent with the direction of these *R* values, we calculate the correlations between them. Due to the fact that the feature weights exhibited in a non-normal distribution, the Spearman's correlation was employed here. Results demonstrated a high correlation of r[female]= 0.8191, p <10^{-30} in females, and r[male]= 0.8513, p <10^{-30} in males (**Figure S4a**). To exclude the influences of feature weights approaching 0, we repeated the above analysis by restricting FCs to those with an absolute weight >10 or >50. Overall, results remain substantially unchanged under the threshold of 10 (r[female]= 0.7955, p <10^{-30}; r[male]= 0.8937, p <10^{-30}; **Figure S4b**) or 50 (r[female]= 0.9029, p <10^{-30}; r[male]= 0.7999, p <10^{-30}; **Figure S4c**). These results suggested that the direction of the identified FC weights also relates to the true direction of the FC-IQ correlations.

Table S3. Consensus functional connections derived from the prediction framework.

Consensus Functional Connectivity for Females					
Node 1	**MNI**	**Node 2**	**MNI**	***r***	***p***
R. paracentral lobule	(10,-34,54)	R. middle occipital gyrus	(34,-86,11)	0.210	0.005
L. paracentral lobule	(-8,-38,58)	R. rostroventral FuG	(33,-15,-34)	0.158	0.037
L. dorsolateral SFG	(-18,24,53)	L. medioventral FuG	(-31,-64,-14)	0.218	0.004
R. rostrodorsal IPL	(47,-35,45)	R. rostral STG	(56,-12,-5)	0.286	1.27×10^{-4}
L. caudal SPL	(-15,-71,52)	L. nucleus accumbens	(-17,3,-9)	0.345	3.21×10^{-6}
R. rostral SPL	(19,-57,65)	R. rostral IFG	(51,36,-1)	-0.222	0.003
L. postcentral gyrus	(-21,-35,68)	R. posterior thalamus	(15,-25,6)	-0.289	1.08×10^{-4}
R. medial orbital gyrus	(6,57,-16)	L. medial thalamus	(-7,-12,5)	0.092	0.227
L. caudal PhG	(28,-8,-33)	L. occipital polar cortex	(-18,-99,2)	-0.256	6.61×10^{-4}
L. caudal PhG	(28,-8,-33)	L. inferior occipital cortex	(-30,-88,-12)	-0.362	9.26×10^{-7}
R. caudal cuneus	(8,-90,12)	R. middle occipital gyrus	(34,-86,11)	-0.304	4.50×10^{-5}
L. STG, TE1.0/1.2	(-50,-11,1)	R. rostral lingual gyrus	(18,-60,-7)	-0.284	1.49×10^{-4}
L. rostral MTG	(-53,2,-30)	R. superior occipital cortex	(29,-75,36)	0.198	0.009
Consensus Functional Connectivity for Males					
Node 1	**MNI**	**Node 2**	**MNI**	***r***	***p***
R. medial SFG	(6,38,35)	R. medial precuneus	(7,-47,58)	0.251	5.48×10^{-4}
L. dorsal IFG	(-46,13,24)	R. occipital thalamus	(13,-27,8)	-0.228	0.002
L. medial orbital gyrus	(-7,54,-7)	L. caudal cingulate gyrus	(-7,-23,41)	0.244	7.73×10^{-4}
L. lower limb PCL	(-8,-38,58)	R. rostral MTG	(51,6,-32)	-0.213	0.004
R. lower limb PCL	(10,-34,54)	R. caudal lingual gyrus	(10,-85,-9)	-0.184	0.012
L. STG	(-54,-32,12)	R. rostrodorsal IPL	(39,-65,44)	0.242	8.86×10^{-4}
R. caudal IPL	(45,-71,20)	R. middle occipital gyrus	(34,-86,11)	-0.306	2.15×10^{-5}
R. occipital polar cortex	(22,-97,4)	R. rostral temporal thalamus	(3,-13,5)	0.233	0.001

Table S4. Predictability of the consensus FCs from the same or opposite gender group

		Consensus FCs of the same gender		Consensus FCs of the opposite gender	
		Males	**Female**	**Males**	**Females**
Intelligence	Information	**0.491 ± 0.010**	**0.503 ± 0.012**	0.122 ± 0.027	0.161 ± 0.024
	Similarity	**0.393 ± 0.017**	**0.505 ± 0.013**	0.054 ± 0.032	0.160 ± 0.027
	Digital span	**0.313 ± 0.018**	**0.516 ± 0.013**	0.072 ± 0.034	0.048 ± 0.033
	Digital symbol	**0.273 ± 0.017**	**0.080 ± 0.027**	0.060 ± 0.028	0.102 ± 0.027
	Picture comprehension	**0.249 ± 0.018**	**0.398 ± 0.020**	0.064 ± 0.033	0.167 ± 0.024
	Block design	**0.452 ± 0.011**	**0.518 ± 0.012**	0.145 ± 0.031	0.028 ± 0.036
	IQ	**0.569 ± 0.009**	**0.706 ± 0.007**	0.060 ± 0.029	0.179 ± 0.027
Temperament	Novelty seeking	0.082 ± 0.027	0.075 ± 0.036	0.101 ± 0.021	0.095 ± 0.031
	Harm avoidance	0.065 ± 0.028	0.052 ± 0.038	0.091 ± 0.027	0.107 ± 0.046
	Reward dependence	0.101 ± 0.039	0.100 ± 0.032	0.065 ± 0.039	0.037 ± 0.027

Prediction results for 10 behavior metrics (7 intelligence and 3 temperament) scores solely based on the consensus FCs from the same or opposite gender group for male and female subjects respectively. Here we ran multiple linear regression with 10-fold cross-validation, in which the consensus FCs were used as regressors and each of the ten behavior metrics was used as the target measure. The process was performed with 100 bootstrapping repetitions with subjects randomly shuffled for each of the 10 behavior metrics. Specifically, the results revealed significant correlations between the predicted and observed IQ scores for males and females. Regarding the 6 intelligence sub-domains, only the digit symbol was not significantly predicted by consensus FCs in females. In both gender groups, no significant correlations were achieved for 3 temperament traits. There were no significant prediction results in males when predicting any of the ten behavior metrics with female-specific consensus FCs, and neither in females similarly. Moreover, similar to the prediction results using whole-brain FCs, females achieved significantly higher accuracies than males for almost all 7 metrics ($p < 0.001$) except the digital symbol.

REFERENCES

Beaty RE, Kenett YN, Christensen AP, Rosenberg MD, Benedek M, Chen Q, Fink A, Qiu J, Kwapil TR, Kane MJ, Silvia PJ. 2018. Robust prediction of individual creative ability from brain functional connectivity. Proc Natl Acad Sci U S A.

Dosenbach NU, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, Nelson SM, Wig GS, Vogel AC, Lessov-Schlaggar CN. 2010. Prediction of individual brain maturity using fMRI. Science 329:1358-1361.

Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT. 2015.

Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosci 18:1664-1671.

First MB, Spitzer, R.L., Gibbon, M. and Williams, J.B. 1995. Structured Clinical Interview for DSM-IV axis I disorders. New York: Biometrics Research Department,New York State Psychiatric Institute.

Freire L, Roche A, Mangin JF. 2002. What is the best similarity measure for motion correction in fMRI time series? IEEE Trans Med Imaging 21:470-484.

Friedman J, Hastie T, Tibshirani R. 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw 33:1-22.

Friston KJ, Ashburner J, Frith CD, Poline JP, Heather JD, Frackowiak RS. 1995. Spatial registration and normalization of images. Hum Brain Mapp 2:165-189.

Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinson M, Consortium WU-MH. 2013. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80:105-124.

Greene AS, Gao S, Scheinost D, Constable RT. 2018. Task-induced brain state manipulation improves prediction of individual traits. Nature communications 9:2807.

Jiang R, Abbott CC, Jiang T, Du Y, Espinoza R, Narr KL, Wade B, Yu Q, Song M, Lin D, Chen J, Jones T, Argyelan M, Petrides G, Sui J, Calhoun VD. 2018. SMRI Biomarkers Predict Electroconvulsive Treatment Outcomes: Accuracy with Independent Data Sets. Neuropsychopharmacology 43:1078-1087.

Karatzoglou A, Smola A, Hornik K, Zeileis A. 2004. kernlab - An S4 Package for Kernel Methods in R. Journal of Statistical Software 11:721-729.

Kaufman S, Rosset S, Perlich C, Stitelman O. 2012. Leakage in Data Mining: Formulation, Detection, and Avoidance. Acm T Knowl Discov D 6.

Liu Z, Zhang J, Xie X, Rolls ET, Sun J, Zhang K, Jiao Z, Chen Q, Zhang J, Qiu J, Feng J. 2018. Neural and genetic determinants of creativity. NeuroImage 174:164-176.

Meng X, Jiang R, Lin D, Bustillo J, Jones T, Chen J, Yu Q, Du Y, Zhang Y, Jiang T, Sui J, Calhoun VD. 2017. Predicting individualized clinical measures by a generalized prediction framework and multimodal fusion of MRI data. Neuroimage 145:218-229.

Reggente N, Moody TD, Morfini F, Sheen C, Rissman J, O'Neill J, Feusner JD. 2018. Multivariate resting-state functional connectivity predicts response to cognitive behavioral therapy in obsessive-compulsive disorder. Proc Natl Acad Sci U S A 115:2222-2227.

Robnik-Sikonja M, Kononenko I. 1997. An adaptation of Relief for attribute estimation in regression. In: Fisher DH, editor. ICML: Morgan Kaufmann. p 296-304.

Robnik-Šikonja M, Kononenko I. 2003. Theoretical and Empirical Analysis of ReliefF and RReliefF. Machine Learning 53:23-69.

Rosenberg MD, Finn ES, Scheinost D, Papademetris X, Shen X, Constable RT, Chun MM. 2016. A neuromarker of sustained attention from whole-brain functional connectivity. Nat Neurosci 19:165-171.

Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, Duff E, Feinberg DA, Griffanti L, Harms MP, Kelly M, Laumann T, Miller KL, Moeller S, Petersen S, Power J, Salimi-Khorshidi G, Snyder AZ, Vu AT, Woolrich MW, Xu J, Yacoub E, Ugurbil K, Van Essen DC, Glasser MF, Consortium WU-MH. 2013. Resting-state fMRI in the Human Connectome Project. Neuroimage 80:144-168.

Steiger JH. 1980. Tests for Comparing Elements of a Correlation Matrix. Psychol Bull 87:245-251.

Stokes ME, Visweswaran S. 2012. Application of a spatially-weighted Relief algorithm for ranking genetic predictors of disease. BioData mining 5:20.

Sui J, He H, Yu Q, Chen J, Rogers J, Pearlson GD, Mayer A, Bustillo J, Canive J, Calhoun VD. 2013. Combination of Resting State fMRI, DTI, and sMRI Data to Discriminate Schizophrenia by N-way

MCCA + jICA. Frontiers in human neuroscience 7:235.
Tibshirani R. 1996. Regression shrinkage and selection via the Lasso. J Roy Stat Soc B Met 58:267-288.
Ugurbil K, Xu J, Auerbach EJ, Moeller S, Vu AT, Duarte-Carvajalino JM, Lenglet C, Wu X, Schmitter S, Van de Moortele PF, Strupp J, Sapiro G, De Martino F, Wang D, Harel N, Garwood M, Chen L, Feinberg DA, Smith SM, Miller KL, Sotiropoulos SN, Jbabdi S, Andersson JL, Behrens TE, Glasser MF, Van Essen DC, Yacoub E, Consortium WU-MH. 2013. Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. Neuroimage 80:80-104.
Urbanowicz RJ, Olson RS, Schmitt P, Meeker M, Moore JH. 2018. Benchmarking relief-based feature selection methods for bioinformatics data mining. Journal of biomedical informatics 85:168-188.
Vatansever D, Menon DK, Manktelow AE, Sahakian BJ, Stamatakis EA. 2015. Default Mode Dynamics for Global Functional Integration. J Neurosci 35:15254-15262.
Yamashita M, Yoshihara Y, Hashimoto R, Yahata N, Ichikawa N, Sakai Y, Yamada T, Matsukawa N, Okada G, Tanaka SC, Kasai K, Kato N, Okamoto Y, Seymour B, Takahashi H, Kawato M, Imamizu H. 2018. A prediction model of working memory across health and psychiatric disease using whole-brain functional connectivity. eLife 7.
Yip SW, Scheinost D, Potenza MN, Carroll KM. 2019. Connectome-Based Prediction of Cocaine Abstinence. American Journal of Psychiatry 176:156-164.
Zou H, Hastie T. 2005. Regularization and variable selection via the elastic net. J R Stat Soc B 67:301-320.
Zuo N, Yang Z, Liu Y, Li J, Jiang T. 2018. Both activated and less-activated regions identified by functional MRI reconfigure to support task executions. Brain and behavior 8:e00893.